

**Purdue University**  
**Purdue e-Pubs**

---

ECE Technical Reports

Electrical and Computer Engineering

---

1-1-2009

# Remote Sensing Methods by Compressive Sensing

Atul Divekar

*Purdue University - Main Campus, [divekar@ecn.purdue.edu](mailto:divekar@ecn.purdue.edu)*

Okan Ersoy

*Purdue University - Main Campus, [ersoy@purdue.edu](mailto:ersoy@purdue.edu)*

Follow this and additional works at: <http://docs.lib.purdue.edu/ecetr>



Part of the [Electrical and Computer Engineering Commons](#)

---

Divekar, Atul and Ersoy, Okan, "Remote Sensing Methods by Compressive Sensing" (2009). *ECE Technical Reports*. Paper 385.  
<http://docs.lib.purdue.edu/ecetr/385>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact [epubs@purdue.edu](mailto:epubs@purdue.edu) for additional information.

# Remote Sensing Methods by Compressive Sensing

---

Atul Divekar, Okan Ersoy

School of Electrical and Computer Engineering,  
465 Northwestern Ave.  
Purdue University,  
West Lafayette, IN 47907-2035

## **ABSTRACT**

Compressive Sensing is a recently developed technique that exploits the sparsity of naturally occurring signals and images to solve inverse problems when the number of samples is less than the size of the original signal. We apply this technique to solve underdetermined inverse problems that commonly occur in remote sensing, including superresolution, image fusion and deconvolution. We use  $l_1$ -minimization to develop algorithms that perform as well as or better than conventional methods for these problems. Our algorithms use a library of samples from similar images or a model for the image to be reconstructed to express the image as a sparse linear combination. A set of feature vectors is generated from the library or basis and is used to find the sparsest linear combination that matches the data using  $l_1$ -minimization.

## Table of Contents

Introduction.....	1
Inverse problems and Regularization.....	1
Compressive Sensing .....	1
Restricted Isometry Property.....	2
Superresolution .....	4
Image Fusion.....	7
A Low-Complexity Modified Brovey Transform with Low Color Distortion .....	11
Comparison of Fusion algorithms .....	13
Deconvolution.....	15
Bibliography .....	16

## Introduction

### Inverse problems and Regularization

Inverse problems are commonly found in image processing, especially when working with remotely sensed images. Inverse problems involve the estimation or reconstruction of a signal given some measurements that are functions of the signal. Typically an inverse problem is characterized by an equation of the form  $y = \Phi x + \eta$ , where  $y$  is a length  $M$  vector of samples obtained from a sensor such as a satellite camera,  $\Phi$  is an  $M \times N$  matrix representing the sensing mechanism and  $x$  is the length  $N$  signal vector that we need to reconstruct. Here  $\eta$  represents noise introduced by the sensing process. Image fusion, super resolution, deconvolution, and classification are examples of problems that can be characterized by this equation.

Many natural signals such as images are sparse in an appropriately chosen domain, i.e. their coefficients with respect to an orthonormal basis such as discrete cosine or wavelet or in an overcomplete dictionary decay rapidly in magnitude (for example, according to a power law).

This means that most of the energy of the signal is contained in a relatively small number of coefficients. As a special case, a signal may have only a few non-zero coefficients.

### Compressive Sensing

Compressive Sensing is a new signal processing topic that exploits the sparsity of coefficients of natural signals for the solution of underdetermined inverse problems.

If a signal is known to have only a few nonzero coefficients (i.e. only some entries of  $c$  are nonzero), is it possible to recover these from a few linear measurements of the signal? For now, assume that the noise  $\eta=0$ . In general, recovering the nonzero coefficients would require one to solve the following problem:

$$\begin{aligned} \min \quad & \|c\|_0 \text{ such that} \\ & y = \Phi \Psi c \end{aligned}$$

Here  $\|c\|_0$  is the number of nonzero coefficients in  $c$ . This is a combinatorial optimization problem and is known to be NP-complete. However, it has been shown in [1] that under certain conditions, the solution to this problem is the same as the solution to the corresponding  $l_1$  minimization problem:

$$\begin{aligned} \min \quad & \|c\|_1 \text{ such that} \\ & y = \Phi \Psi c \end{aligned} \quad \dots (1)$$

Here  $\|c\|_1 = \sum_{i=1}^N |c_i|$  is the  $l_1$ -norm of  $c$ . This is a convex function of  $c$ .

This is a convex optimization problem that can be solved tractably by linear programming methods. For example, interior point methods can solve it in polynomial time.

Greedy algorithms such as Orthogonal Matching Pursuit (OMP) [2] have been developed to solve the same problem with weaker guarantees.

### Restricted Isometry Property

In [1], Tao et. al. define a sufficient condition called the Restricted Isometry Property (RIP) that guarantees perfect recovery of all the nonzero coefficients. This property is elaborated below.

Let  $U = \Phi\Psi$ . The matrix  $U$  (size  $M \times N$ ) satisfies the RIP with parameters  $(m, \varepsilon)$  for  $\varepsilon \in (0, 1)$  if we have

$$(1 - \varepsilon) \|c\|_2^2 \leq \|Uc\|_2^2 \leq (1 + \varepsilon) \|c\|_2^2 \quad \text{for all } m\text{-sparse vectors}$$

This means that for each set of  $m$  columns of  $U$ , the eigenvalues lie between  $(1 - \varepsilon)$  and  $(1 + \varepsilon)$ . Intuitively, the Restricted Isometry Property states that every set of  $m$  columns of  $U$  behaves approximately as if they were orthonormal. Since the initial result in [1], several improvements have been made on the parameters  $(m, \varepsilon)$  for which the signal recovery is provably guaranteed. Currently the best result from [3] states that perfect recovery is possible (if  $\eta=0$ ) for  $m = M/2$  if  $\varepsilon < \sqrt{2} - 1$ . In words, if the number of linear combinations (samples)  $M$  is twice the number of non-zero components of  $c$ , *whatever their locations in  $c$  might be*, and the RIP is satisfied with  $\varepsilon < 0.41$ , then  $l_1$ -minimization can recover the coefficient vector  $c$  exactly.

The results in [3] go further. Again assume that  $\eta=0$ , i.e. the noiseless case. Most real signals are not exactly sparse with respect to any basis  $\Psi$ , but have rapidly decaying coefficient values (for example, proportional to  $i^{-s}$ , where  $i$  is the index of the coefficient when sorted in descending order of magnitude, and  $s > 0$  is the power of decay). Suppose that we set to zero all but the largest  $m$  coefficients by magnitude in  $c$  and let  $c_m$  denote this approximation to  $c$ . Suppose that the solution to the  $l_1$ -minimization program is  $c^*$ . Then [3] shows that

$$\|c^* - c\|_1 \leq C_0 \|c - c_m\|_1$$

and

$$\|c^* - c\|_2 \leq C_0 m^{-1/2} \|c - c_m\|_1$$

Here  $C_0$  is a small constant. (For  $\varepsilon=0.2$ ,  $C_0=4.2$  is sufficient).

A similar result is possible for the noisy case. Suppose that the noise magnitude present satisfies  $\|\eta\|_2 < \delta$ . Now we solve the convex optimization problem

$$\begin{aligned} & \min \|c\|_1 \\ & \text{subject to } \|y - Uc\|_2 < \delta \end{aligned}$$

If  $\varepsilon < \sqrt{2} - 1$ , the solution  $c^*$  to this program obeys  $\|c^* - c\|_2 \leq C_0 m^{-1/2} \|c - c_m\|_1 + C_1 \delta$ .

(For  $\varepsilon=0.2$ , we have  $C_1=8.5$  and  $C_0$  same as before).

We use the results of Compressive Sensing to develop novel algorithms for Image Superresolution and Deconvolution.

## Superresolution

Given a blurred or low resolution image (or possibly more than one image) we wish to reconstruct a higher resolution image. Here we have  $y = \Phi x + \eta$  where  $x$  is the vectorized  $N \times N$  high resolution image to be reconstructed and  $y$  is the low-resolution  $N/2 \times N/2$  image with  $\Phi$  the filtering matrix. We may try to obtain the solution as

$$\text{find } x \text{ to min } \|y - \Phi x\|_2^2$$

However this has no unique solution. The problem is ill-posed: knowledge of the forward model does not provide a single solution by inversion. A common approach [4] is to regularize the problem by adding a constraint that reflects *a priori* knowledge about the domain of image  $x$ . This converts the problem into a fully determined problem with a unique solution. Commonly, the smoothness of the image –a property of most natural images- is used as a constraint. For example, we modify the solution to

$$\text{find } x \text{ to min } \|y - \Phi x\|_2^2 + \lambda \|Dx\|_2^2$$

Here  $D$  is a high pass filter such as the Laplacian kernel matrix. The second term penalizes the differences between neighboring pixels, and  $\lambda$  is the Lagrange multiplier that determines the relative significance of the first and second terms. algorithm uses a dictionary  $D_H$  of  $4 \times 4$  pixel patches taken from high resolution training images that have the same statistical properties as the image to be reconstructed. Each patch has its mean subtracted out. For each patch in  $D_H$  we produce a low-resolution “sample” patch by blurring with the same operator  $\Phi$  used in the forward model. The dictionary  $D_L$  of low-resolution patches is used for  $l_1$  minimization to reconstruct each  $4 \times 4$  high resolution patch. To ensure continuity of features in the reconstructed image, we use overlapped patches with the left and upper 1-pixel strips of the current patch taken from the already reconstructed left and upper neighbor patches. This provides 7 more “samples” to add to  $D_L$ . Thus the basic algorithm is

From the training images

- 1) Obtain a  $4 \times 4$  size patch dictionary  $D_H$  (size  $16 \times K$ , where  $K$  is the number of samples).
- 2) For each patch in  $D_H$  construct a sample vector that has four  $2 \times 2$  pixel means and the same 7 samples as the left and top 1-pixel strips of the high resolution patch. This gives the low resolution sample dictionary  $D_L$ . This has size  $11 \times K$ . Find the means  $mD_H$  and  $mD_L$  of  $D_H$  and  $D_L$  respectively. Set  $D_H(:, k) \leftarrow D_H(:, k) - mD_H$  and  $D_L(:, k) \leftarrow D_L(:, k) - mD_L$  for each column  $k$ .



- 3) Normalize each column of  $D_L$  to have unit norm. Store the norms in vector  $n$ .

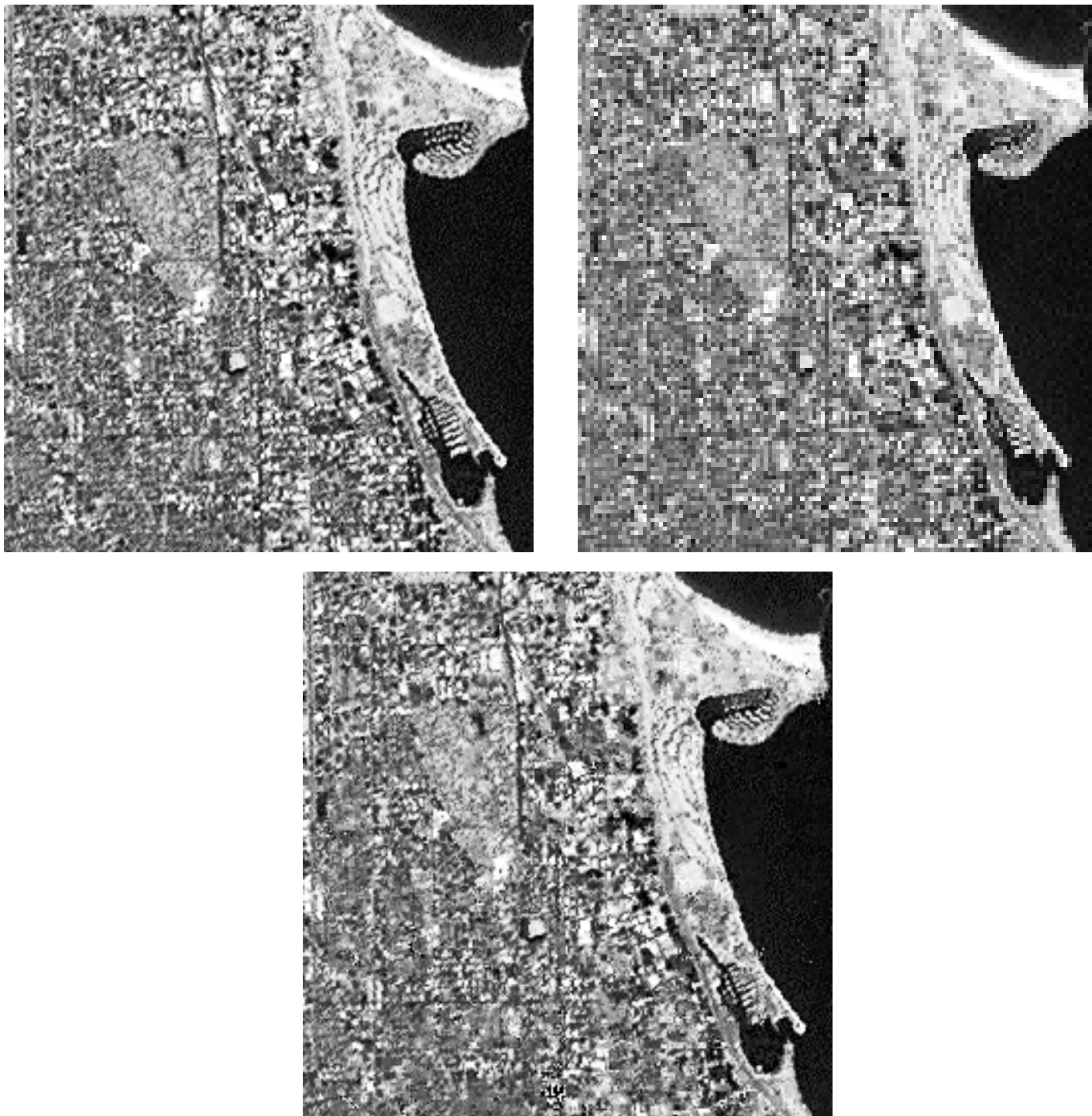
To reconstruct the high resolution image:

For each 4x4 patch in raster order and 1 pixel overlap with previously reconstructed patches,

- 1) Use low resolution pixels (2x2) and samples from left and upper reconstructed patches to construct length 11 vector  $y$ . (For the top and leftmost rows of patches, we use an estimate of the top and/or left pixel edges using a standard method such as Brovey). Set  $y \leftarrow y - mD_L$ .
- 2) Solve 
$$\begin{array}{ll} \min & \|a\|_1 \\ \text{such that} & y = D_L a \end{array}$$
- 3) Normalize  $a \leftarrow a / n$ . The estimate for the high resolution patch is  $\hat{x} = D_H a + mD_H$ .

The performance of this algorithm depends on the similarity of the patches in the dictionary to the actual patterns present. (If the exact pattern is present in the dictionary it will always be recovered provided that every pair of columns of  $D_L$  is linearly independent). For the Landsat training images we used, about 10000 patches are sufficient to recover high resolution image with sufficient visual quality.

Since each low resolution patch is obtained by a blurring operation, it is possible for two high resolution patches to map to a single (or negligibly different) low resolution vector. In this case the  $l_1$ -minimization algorithm can pick the wrong signal as the high resolution reconstruction. However, providing the top and left 1-pixel strips seems to be sufficient to obtain good separation between otherwise similar low-resolution patches. If the mean is subtracted from  $D_L$  we get Gaussian statistics. In [1] it is proven that such a Gaussian matrix has the Restricted Isometry Property with almost always, and hence  $l_1$  minimization will recover the correct linear combination to reconstruct the high resolution image.



**Figure 1: Top Left: High resolution image, Top Right: Blurred image , Bottom: Superresolved image using Compressive Sensing and the *single* blurred image**

## Image Fusion

Image fusion refers to the process of combining several images (possibly from different sensors, viewpoints and resolutions) of a given scene to extract information that is not apparent from any single image alone. This has several applications in remote sensing, medical imaging, video surveillance, etc. Common methods for image fusion include the Intensity-Hue-Saturation (IHS) [5], Brovey transform [6], Principal Component Analysis [7] and wavelet based methods (such as with the *a trous* algorithm) [8] [9]. We consider a common situation occurring in remote sensing. Satellites such as Landsat 7 carry a multispectral sensor and a panchromatic camera. Each multispectral image has a resolution of 30m while the panchromatic image has a resolution of 15m. Different kinds of earth cover emit radiations in different bands while the panchromatic camera provides twice the spatial detail as the multispectral sensor. We use images from three bands and the panchromatic camera. If the high resolution panchromatic image is size  $N \times N$ , each band image is  $N/2 \times N/2$ . We wish to reconstruct three high resolution images each of size  $N \times N$ .

The fused images must not distort the information carried by the original image. This means that the fused images must match the spatial resolution of the panchromatic image while maintaining the spectral properties of the low resolution MS bands.

We suggest two algorithms that use Compressive Sensing to fuse Landsat images. In each algorithm we use a model of the expected fused image to generate a library of possible fused images. The first uses images from the standard PCA method as a starting point while the second uses segmentation.

### A) Algorithm I:

We know that the PCA method of fusion produces good spatial details, but suffers from color distortion. So we use the PCA output to generate the new samples.

The algorithm is:

- 1) Generate the fused images according to the PCA algorithm. Call the fused bands  $b_{p1}$ ,  $b_{p2}$ , and  $b_{p3}$ .

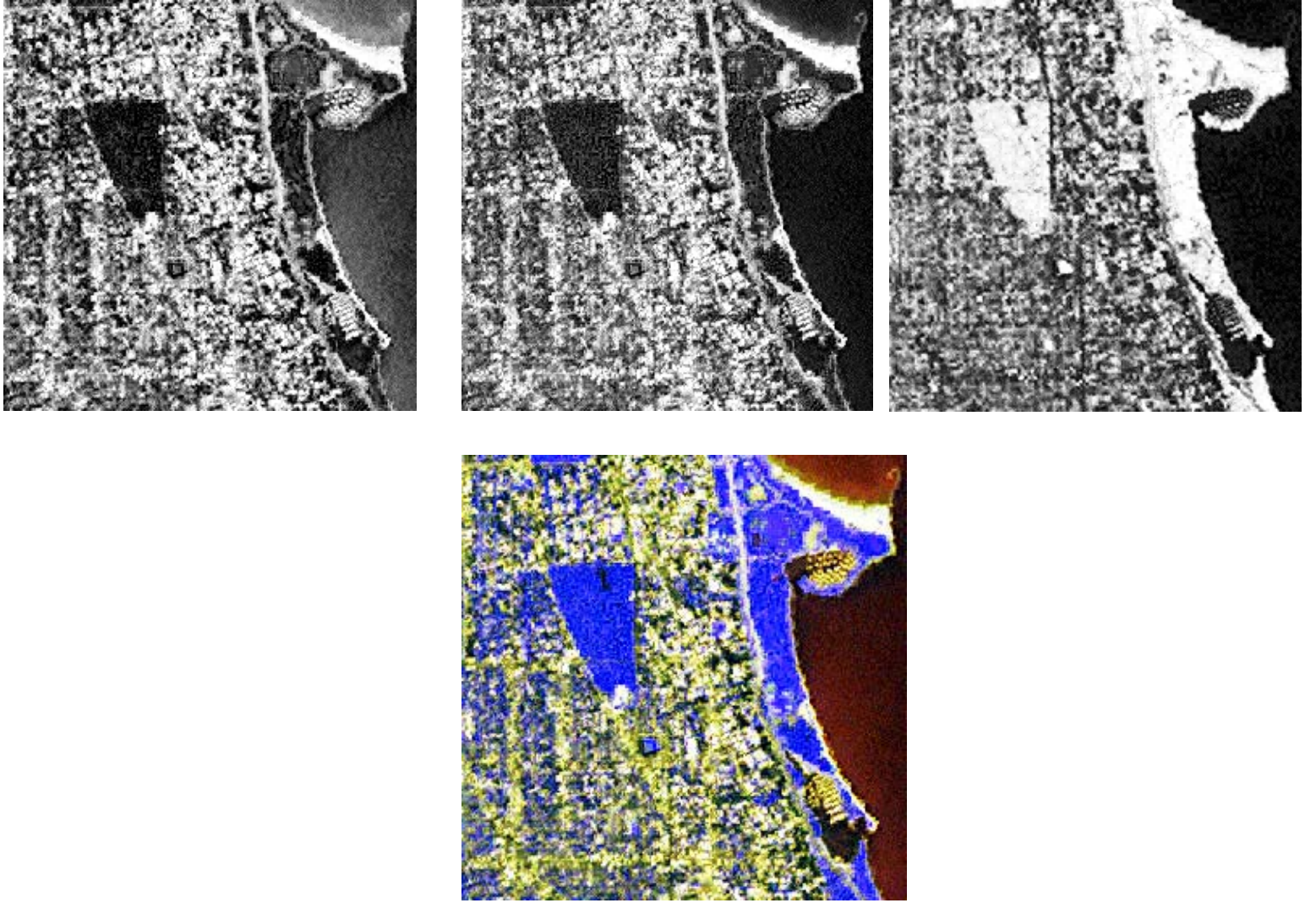
For each  $K \times K$  size block to fuse

- 1) Generate each sample in  $D_H$ :  
For each  $2 \times 2$  block in each fused band  $b_{pi}$ , subtract out the mean  $\mu$  and add a random number drawn from a Gaussian distribution. Store the means  $\mu$  in a vector  $\bar{\mu}$ .

Place this random number as the 2x2 mean value in  $D_L$ .

- 2) Find the means  $mD_H$  of  $D_H$  and  $mD_L$  of  $D_L$ . Subtract these out of each term in  $D_H$  and  $D_L$  respectively.
- 3) From the given low resolution MS images, get blocks to form  $y$ , a  $3 \times K \times K/4$  length vector. Set  $y \leftarrow y - mD_L$ .
- 4) Let the column norms of  $D_L$  be stored in vector  $n$ . Normalize each column of  $D_L$  to unit norm.
- 5) Solve  $\min_{\text{such that } y=D_L a} \|a\|_1$ . Set  $a \leftarrow a/n$ . The reconstruction is  $\hat{x} = mD_H + \bar{\mu} + D_H a$ .

The result is shown in Figure 2.



**Figure 2: Fusion by Compressive Sensing I**

B) Algorithm 2:

We propose another image fusion algorithm using segmentation. This is based on the fact that the panchromatic image already contains the high resolution detail, and that if the contiguity of each object in the pan image is maintained in the fused image, the details are likely to be preserved. We seek to build a dictionary that provides possible candidates for the fused image patches *given the panchromatic patch at the corresponding position*. To preserve the fine detail in the panchromatic image we must approximately *maintain the difference in mean values* of adjacent features in the pan image.

We also want them to match the low-resolution image block in the multispectral image.

We utilize the fact that the high to low resolution sub sampling operation is carried out on blocks of size 2x2 pixels. However, object boundaries are not well aligned with 2x2 block boundaries. In a sense, object boundaries are *incoherent* to the sampling grid. Also, objects are restricted to a small subset of the image.

For each KxK patch  $p$  the panchromatic image,

- 1) Let  $\mu$  be its mean and let  $\mu_1, \mu_2$  and  $\mu_3$  be the means of the 2x2 blocks in each low resolution band image that (when interpolated) cover the same spatial area as  $p$ . Form  $p_i = p - \mu + \mu_i$  ( $i=1, 2, 3$ ).

For each band  $i$ ,

- 2) Let  $p_{dn}$  be the zero-mean blurred version of the pan block  $p$  and let  $b_{iz}$  be the low resolution zero mean MS band block. Find  $\rho_i = \frac{\langle p_{dn}, b_{iz} \rangle}{\|p_{dn}\| \|b_{iz}\|}$  Form  $z = \rho^2 p_{dn} + (1 - \rho^2) b_{iz}$  and segment  $z$  by k-means. Find the adjacency matrix.

- 3) Create dictionary  $D_H$  of size  $K \times K$  by  $S$  where  $S$  is the number of samples.

- 4) To create a sample image  $V$  to form a column in  $D_H$

For each pair of adjacent regions  $i$  and  $j$  with supports  $I_i$  and  $I_j$ ,

- a) Generate a random number  $r$  from an  $N(0, 1)$  distribution.
- b)  $V(I_i) \leftarrow V(I_i) + r$ ;  $V(I_j) \leftarrow V(I_j) - r$

- 5) Obtain the low-resolution dictionary  $D_L$  of size  $K \times K/4$  by  $S$  by blurring each image in  $D_H$ . Normalize  $D_L$  to get unit norm columns. Let the vector of norms be  $n$ .

- 6) Let  $y$  be the vector of 2x2 block means from the low resolution MS band. Set  $y \leftarrow y - mD_L$ . Here  $mD_L$  is the low resolution version of  $p_k$

- 7) Solve  $\min_{\text{such that } y = D_L a} \|a\|_1$ . Normalize  $a \leftarrow a / n$ . Reconstruct  $\hat{x} = p_k + D_H a$ .

The result of this algorithm is shown in Figure 3.



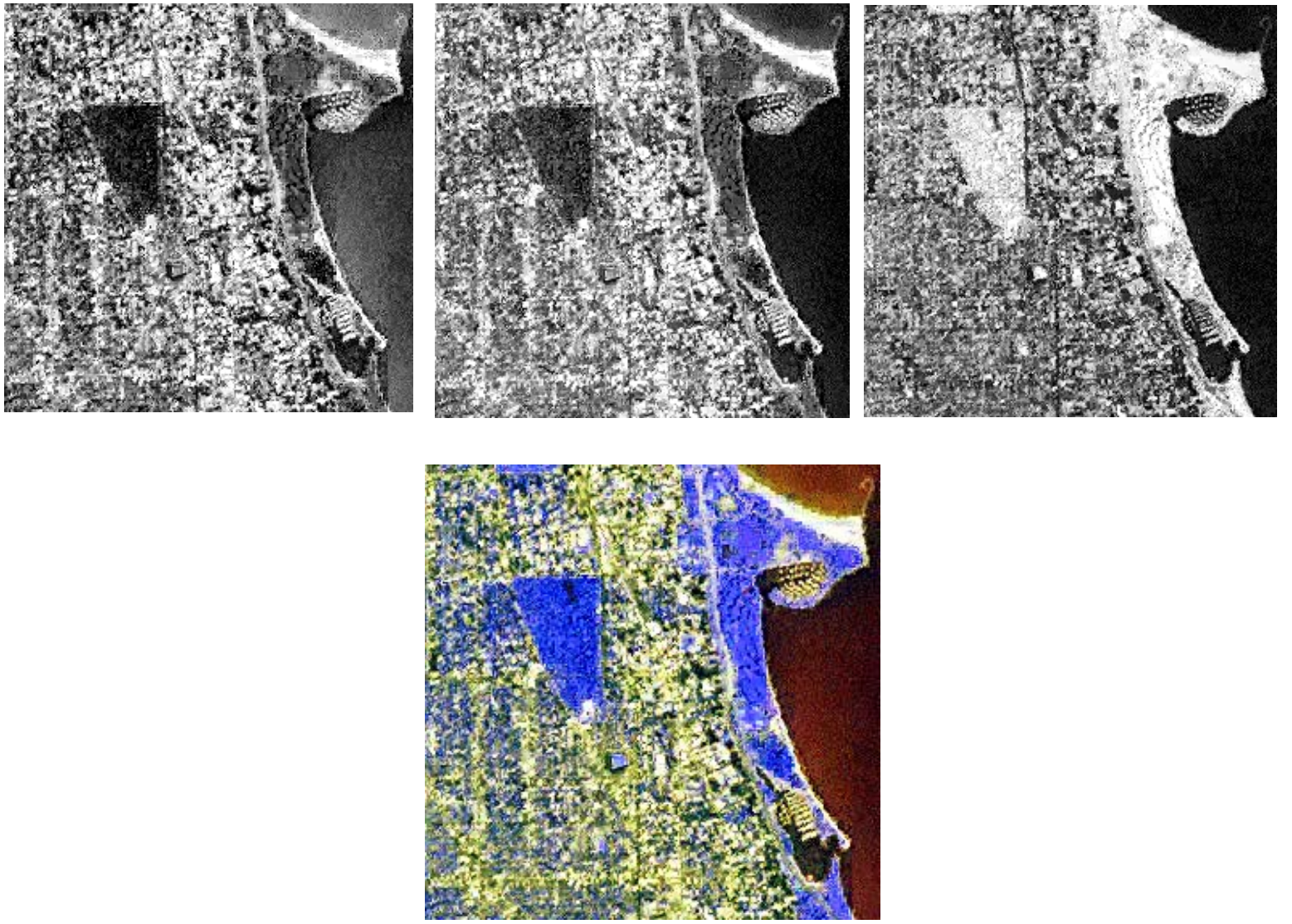


Figure 3: Result of Image Fusion by Compressive sensing II- segmentation

### A Low-Complexity Modified Brovey Transform with Low Color Distortion

The standard Brovey transform is defined in the following way. Let  $p$  be a  $2 \times 2$  pixel block from the panchromatic image and let  $b_1$ ,  $b_2$  and  $b_3$ . Then the  $2 \times 2$  blocks in each fused image corresponding to the location of  $p$  are defined as  $b_i^f = \frac{b_i}{\sum_{j=1..3} b_j} \cdot p$ . For this block, let  $\mu$  be the mean

and  $p_d = p - \mu$ . We can write  $b_i^f = b_i \left( \frac{p_d + \mu}{\sum_{k=1..3} b_k} \right)$ . Each fused pixel value is a product of  $b_i$  and a

multiple which has the same value over all the bands at a particular pixel location. This allows the value of spectral distortion measures such as the Spectral Angle Mapper to be 0 degrees.

However this multiple changes according to pixel location and this causes serious color distortion as seen in **Error! Reference source not found.**

To reduce this distortion we define the following modified transform:

For each 2x2 block of the pan image we define the 2x2 fused image patches as

$$b_i^f = b_i \left(1 + \frac{p_d}{\sum_{j=1,3} b_j}\right).$$

This keeps the multiple approximately identical over different pixel locations

and reduces spectral distortion. The result of this algorithm is shown in

The effect of reduced spectral distortion is seen in the images as well as in the low value of the ERGAS measure relative to the standard Brovey transform as seen in Table 1.

This transform is the similar to the “mean” value  $mD_H$  obtained in the Compressive Sensing algorithm above with a block size of 2x2. Hence it can be regarded as a low-cost fusion algorithm without the additional correction provided by the Compressive Sensing algorithm. This is a modified Brovey transform that has almost zero spectral distortion and low algorithmic complexity at the cost of some spatial detail.

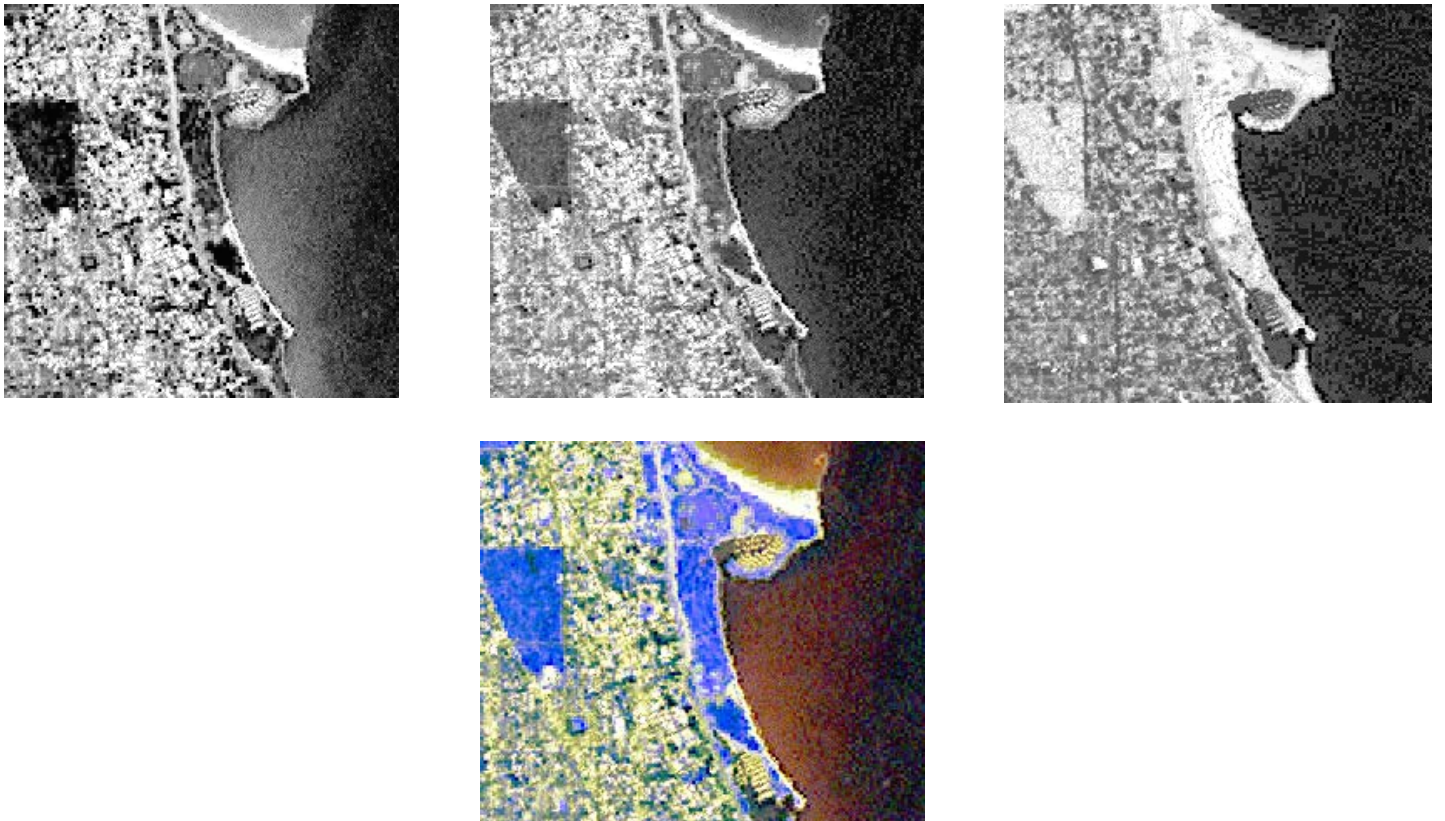


Figure 4: Result of image fusion by low-complexity modified Brovey Transform



## Comparison of Fusion algorithms

We compare the above methods using the commonly used metrics: Correlation Coefficient (CC), Spectral Angle Mapper (SAM) and ERGAS [10] . These are defined below.

Let  $p$  be the  $K \times K$  panchromatic image,  $b_{1d}, b_{2d}, b_{3d}$  be the  $K/2 \times K/2$  low resolution multispectral band images,  $b_{1i}, b_{2i}, b_{3i}$  be the interpolated band images, and  $\hat{b}_1, \hat{b}_2, \hat{b}_3$  be the fused  $K \times K$  images.

$$\text{Correlation Coefficient } CC_i \text{ for band } i = \frac{(p - \mu_p)^T (\hat{b}_i - \mu_{\hat{b}_i})}{\|p - \mu_p\|_2 \|\hat{b}_i - \mu_{\hat{b}_i}\|_2}$$

Here  $\mu_0$  is the mean of the corresponding band. Higher values of Correlation Coefficient indicate better spatial fidelity with the panchromatic image.

Spectral Angle Mapper: Let  $\hat{v}_j = \begin{bmatrix} \hat{b}_{1_j} \\ \hat{b}_{2_j} \\ \hat{b}_{3_j} \end{bmatrix}$  be the pixel vector in the fused bands at location  $i$ , and let

$v_j = \begin{bmatrix} b_{1_{i_j}} \\ b_{2_{i_j}} \\ b_{3_{i_j}} \end{bmatrix}$  be the pixel vector from the interpolated bands at location  $i$ . Then we define the

Spectral Angle Mapper value as  $\frac{1}{K^2} \sum_{j=1}^{K^2} \arccos\left(\frac{\langle v_j, \hat{v}_j \rangle}{\|v_j\| \|\hat{v}_j\|}\right)$ . It measures the average correlation

between color vectors in the original and fused band images. A low value indicates good spectral fidelity. However note that this value is 0 if each  $\hat{v}_j$  is a scalar multiple of the corresponding  $v_j$ .

This happens with the standard Brovey transform algorithm. However, this *does not* lead to good visual color fidelity since the constant changes from one  $2 \times 2$  block to the next.

ERGAS is a frequently used quality measure defined in [11]. It stands for “*erreur relative globale adimensionnelle de synthese*” which means relative dimensionless global error in synthesis. It measures properties that the synthetic (fused) image should try to achieve:

- 1) Each high resolution image  $\hat{b}_j$  on being degraded (blurred) to low resolution should be as identical as possible to the given low resolution image  $b_{jd}$ .
- 2) Each high resolution image  $\hat{b}_j$  should be as similar to the images that the multispectral sensor would capture if it worked at the higher resolution.
- 3) The set of high resolution images  $\hat{b}_1, \hat{b}_2, \hat{b}_3$  should be as identical as possible to the multispectral set of images that the corresponding sensor would observe with the highest spatial resolution  $h$ .

ERGAS is defined as  $100 \frac{h}{l} \sqrt{\frac{1}{N} \sum_{j=1..3} \frac{RMSE_i^2}{\mu_i^2}}$ ; here  $RMSE_i$  is the root mean square error

between the fused and interpolated low-resolution image of band  $i$ ,  $\mu_i$  is the mean of the band  $i$  image and  $h/l$  is the ratio of the size of low resolution to high resolution pixels.

(It is  $\frac{1}{2}$  if each low resolution image pixel is the mean of a 2x2 block of high resolution pixels).

A low value of ERGAS indicates good fidelity to the data.

Method	SAM (deg)	CC			ERGAS
Standard Brovey	0	0.79	0.87	0.74	39.4
Principal Components	9.59	0.91	0.93	0.77	21.43
Compressive sensing-I	5.84	0.54	0.58	0.61	11.62
Compressive sensing-Seg	6.85	0.68	0.77	0.93	13.5
Low Cost Brovey	0	0.48	0.60	0.79	4.84

**Table 1: Comparison of Image Fusion Methods**

## Deconvolution

Images recorded by a sensor are usually blurred due to convolution effects. This is due to the sensor's Point Spread Function (PSF). Thus the signal recorded is  $y = x * h$ , where  $x$  is the original signal and  $h$  represents the Point Spread Function, which is the signal received if  $x$  is assumed to be a point source. If the PSF is known, algorithms such as Lucy-Richardson deconvolution [12] [13] or Wiener deconvolution [14] may be used. If the PSF is unknown it must be estimated either indirectly or simultaneously with the image restoration. Iterative techniques such as the Expectation-Maximization algorithm may be used [15].

When  $H$  (the blurring matrix) is known the algorithm is similar to the superresolution algorithm discussed earlier. We use  $4 \times 4$  sample patches in  $D_H$  and their corresponding  $2 \times 2$  block means in  $D_L$ . To adequately distinguish between patches with the same mean values we add the top and left 7 pixels from the high resolution patch.



Figure 5: Left: Original Image, Middle: Image Blurred by Point Spread Function Right: Image Recovered by Compressive Sensing

## Bibliography

- [1] E. Candes and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. on Information Theory*, vol. 52, no. 12, pp. 5406-5425, Dec. 2006.
- [2] T. Joel and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. on Information Theory*, vol. 53, no. 12, pp. 4655-4666, Dec. 2007.
- [3] E. Candes, "The restricted isometry property and its implications for compressed sensing," *Compte Rendus de l'Academie des Sciences, Paris*, vol. Series I, 346, pp. 589-592, 2008.
- [4] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *International Journal of Imaging Systems and Technology*, vol. 14, no. 2, pp. 47-57, 2004.
- [5] T. Tu, S. Su, S. Shyu, and P. Huang, "A new look at IHS-like image fusion methods," *Information Fusion*, vol. 2, no. 3, pp. 177-186, Sep. 2001.
- [6] a. R. Gillespie, A. Kahle, and R. Walker, "Color enhancement of highly correlated images-II. Channel ratio and chromaticity transformation techniques," *Remote Sensing of Environment*, vol. 22, pp. 343-365, 1987.
- [7] P. S. Chavez, S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data :Landsat TM and SPOT Panchromatic," *Photogrammetric Engineering and Remote Sensing*, vol. 57, no. 3, pp. 311-316, 1985.
- [8] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing*, vol. 57, pp. 234-245, 1995.
- [9] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-Driven Fusion of High Spatial and Spectral IResolution Images Based on Oversampled Multiresolution Analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 10, pp. 3230-3239, Oct. 2002.
- [10] Q. Du, N. Younan, R. King, and V. Shah, "On the Performance Evaluation of Pan-Sharpening Techniques," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 4, pp. 518-522, Oct. 2007.
- [11] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion ?," in

*Proc. Third Conference on "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", Sophia Antipolis, France, 2000, pp. 99-103.*

- [12] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astronomical Journal*, vol. 79, no. 6, pp. 745-754, 1974.
- [13] W. H. Richardson, "Bayesian based Iterative Method of Image Restoration," *Journal of the Optical Society of America*, vol. 62, no. 1, pp. 55-59, 1972.
- [14] R. Gonzalez, R. Woods, and S. Eddins, *Digital Image Processing Using Matlab*. Prentice-Hall, 2003.
- [15] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Stat. Soc Ser. B.*, vol. 39, no. 1, pp. 1-38, 1977.